Published online 21 December 2010 *Nucleic Acids Research, 2011, Vol. 39, No. 8* ***3079–3092***
doi:10.1093/nar/gkq1272

Two groups of phenylalanine biosynthetic operon leader peptides genes: a high level of apparently incidental frameshifting in decoding *Escherichia coli pheL*

Olga L. Gurvich[1], S. Joakim Näsvall[2], Pavel V. Baranov[3], Glenn R. Björk[2] and John F. Atkins[1,4,*]

[1]Department of Human Genetics, University of Utah, Salt Lake City, UT, 84112-5330, USA, [2]Department of Molecular Biology, Umeå University, S-901 87 Umeå, Sweden, [3]Department of Biochemistry and [4]BioSciences Institute, University College Cork, Cork, Ireland

Received July 5, 2010; Revised and Accepted November 23, 2010

ABSTRACT

The bacterial *pheL* gene encodes the leader peptide for the phenylalanine biosynthetic operon. Translation of *pheL* mRNA controls transcription attenuation and, consequently, expression of the downstream *pheA* gene. Fifty-three unique *pheL* genes have been identified in sequenced genomes of the gamma subdivision. There are two groups of *pheL* genes, both of which are short and contain a run(s) of phenylalanine codons at an internal position. One group is somewhat diverse and features different termination and 5′-flanking codons. The other group, mostly restricted to Enterobacteria and including *Escherichia coli pheL*, has a conserved nucleotide sequence that ends with UUC_CCC_UGA. When these three codons in *E. coli pheL* mRNA are in the ribosomal E-, P- and A-sites, there is an unusually high level, 15%, of +1 ribosomal frameshifting due to features of the nascent peptide sequence that include the penultimate phenylalanine. This level increases to 60% with a natural, heterologous, nascent peptide stimulator. Nevertheless, studies with different tRNA^{Pro} mutants in *Salmonella enterica* suggest that frameshifting at the end of *pheL* does not influence expression of the downstream *pheA*. This finding of incidental, rather than utilized, frameshifting is cautionary for other studies of programmed frameshifting.

INTRODUCTION

The shift-prone sequences at which mRNA:tRNA realignment occurs during programmed frameshifting, are generally avoided in highly expressed genes except where the resultant frameshifting is utilized for gene expression. In poorly expressed genes, where deleterious effects at the protein product or mRNA structural/stability levels are minimal, such sequences do not seem to be rare—at least as extrapolated from the occurrences in *Escherichia coli* of the −1 and +1 shift-prone sequences, A_AAA_AAG, CCC_UGA, AGA_AGA and AGG_AGG (1,2). However, when specific frameshifting is selected for gene expression, there are generally stimulatory signals that elevate the level of frameshifting at the shift site. These stimulatory, or recoding signals, are often particular mRNA sequence 3′ of the shift site that commonly form certain mRNA structures, which influence the ribosome centred on the shift site. Other characterized frameshift stimulatory signals are specific 5′-mRNA sequences. Bioinformatic evidence indicates that particular nascent peptide sequences can promote utilized frameshifting (3). Furthermore, a nascent peptide stimulator for the programmed bypassing of 50 nt in decoding phage T4 gene 60 acts by causing peptidyl-tRNA codon:anticodon dissociation. Not surprisingly this nascent peptide can induce frameshifting on synthetic constructs (4).

Specific nascent peptide sequences that act within the exit tunnel of the ribosome can trigger changes within the peptidyl transferase centre (PTC) and/or induce ribosomal pausing. Mediation of these effects, in some cases,

*To whom correspondence should be addressed. Tel: 801 585 3434; Fax +353 238 855 147; Email: atkins@genetics.utah.edu
Present addresses:
Olga L. Gurvich, Cork Cancer Research Centre, BioSciences Institute, University College Cork, Cork, Ireland.
S. Joakim Näsvall, Department of Medical Biochemistry and Microbiology, Uppsala University, Box 582, S-75123, Uppsala, Sweden.

involves exogenous factors. Inducible expression of erythromycin-resistance genes relies on ribosomal stalling, which is dependent on drug binding in the exit tunnel (5); translational arrest during SecM synthesis, which regulates translation of the co-transcribed SecA, is responsive to the secretion status of the cell (6); ribosomal stalling during translation of TnaC, which regulates expression of tryptophanase operon, is sensitive to the level of free tryptophan (7,8). The mechanism of action of several nascent peptide signals, including that of SecM, has been studied in detail (9–12).

Stop codons are slow to decode and UGA is, overall, the least efficient terminator. The extent of the pause at a termination codon is influenced by the identity of the 3′-adjacent base (13) and when the stop codon is preceded by a proline codon, its decoding is especially slow, perhaps because the C-terminal Pro residue facilitates *trans*- to *cis*-isomerization of the final peptide bond, thereby inhibiting termination (14). A C-terminal proline has a variety of translational consequences (15–17); for instance, it is a key feature of the TnaC nascent peptide-mediated effect (18). Additionally, it can facilitate frameshifting. When a slow-to-decode stop codon in the ribosomal A-site is paired with a P-site peptidyl-tRNA that has good potential for re-pairing to mRNA in an overlapping frame, the combination is especially shift-prone, hence the term 'shifty stop' (19,20). Thus, not surprisingly, the sequence CCC_UGA is especially +1 shift-prone (21–23). Frameshifting at CCC_UGA is utilized for expression of antizyme in some eukaryotes (24) and in the expression of the gene for a major tail component in several *Listeria* phages (25,26). Codons other than UGA are also used in the A-site; some phages use +1 frameshifting at the CCC_UAA, i.e. the *Lactobacillus* phage Q54 (27) and *Bacillus* phage SPP1 (28), while in a *Listeria* phage (26) and the *E. coli cheA* gene (29), the slow-to-decode UGA stop codon is substituted by a sense codon whose tRNA is severely limiting.

The frameshifting that occurs at one of the most shifty −1 sites in *E. coli*, A_AAA_AAG (30,31), involves a weak interaction of anticodon base 34 of the plentiful sole $tRNA^{Lys}$ with the third codon bases in the A-site (32,33). More striking is the importance of lack of full Watson–Crick codon complementarity of the P-site tRNA during *Saccharomyces cerevisiae* Ty3 frameshifting (34,35). The cognate tRNA for CCC (36,37) is also often rare, even in *E. coli* and *Salmonella* that do not have a low GC content. To a small extent, this increases the chance of acceptance at CCC-containing ribosomal A-sites, of a more abundant near-cognate proline tRNA with elevated frameshifting consequences. Mutants of $tRNA^{Pro}$ selected for enhanced frameshifting at CCC led to the isolation of partially debilitating mutants of the CCC-decoding isoacceptor that substantially increased the chance of CCC being decoded by a near-cognate tRNA. After its transfer to the P-site, a near-cognate tRNA is more prone than a cognate tRNA to dissociate from mRNA and realign before re-pairing to mRNA (38,39). Some of these mutants are important for the present work.

Of the 19 *E. coli* genes that terminate with CCC_UGA, one, *pheL*, shows a dramatically higher level of frameshifting, 15% in a *relA* strain, than any of the others (1). Although the level of frameshifting is slightly higher in *relA* strains than in WT stringent cells, the effect is small (2,40). This frameshift efficiency is much more than can be accounted for by just CCC_UGA, which is generally not more than 2% (1,21).

pheL is co-transcribed with the 3′ *pheA* whose product catalyses the first two steps of phenylalanine biosynthesis from chorismate (41). Translation of the short *pheL* leader peptide-encoding sequence, in particular the run(s) of phenylalanine codons within it, governs transcription attenuation and, consequently, expression of the downstream *pheA*. When Phe is limiting, the attenuation mechanism (42–44), Figure 1, results in greatly elevated *pheA* mRNA synthesis and expression. However, there is a basal level of expression even when cells are replete for Phe. With abundant Phe, this basal level is ~10% of the level under Phe starvation conditions. The basal level is dependent on the efficiency of ribosome release from the *pheL* UGA stop codon—release factor mutants can decrease the basal level of *pheA* expression (45). The possibility that frameshifting at CCC_UGA may influence basal expression of *pheA* has not been addressed and is unknown in biosynthetic operon expression.

Frameshifting at the CCC_UGA of *pheL* gene results in the synthesis of an extra product. This product can potentially have a separate function. However, since frameshifting allows ribosomes to travel past the stop codon of *pheL* and interfere with formation of secondary structures, it could influence attenuation of transcription and *pheA* expression. If so, it would be the first case where the significance of frameshifting is in mediating ribosomal progression to a novel location instead of in synthesizing an extra protein product (46). [The only specific suggestion for such a concept (47) has not been experimentally investigated.]

MATERIALS AND METHODS

Plasmid constructions

The *pheL* and *pdxH* gene sequences were amplified by PCR from *E. coli* genomic DNA or from previous constructs (1) using primers with appropriate overhanging restriction sites. All subsequent nucleotide changes were introduced using PCR with complementary oligonucleotides carrying appropriate changes as primers. All plasmid constructions were confirmed by DNA sequencing.

For analysis of frameshift stimulators, sequences were cloned into the BamHI/EcoRI sites of pGHM57 (4) in-between GST and maltose binding protein (MBP), so that MBP is in the +1 frame relative to GST. Construction of HLLH, HLH5 and HLH3 was performed by a two-step PCR. In the first step, DNA sequences corresponding to *pheL*, 5′ and 3′ of *pdxH*, were amplified separately. To amplify the *pheL* sequence, the *pheL*/*pdxH* chimera primers HLa(GGAAGATTGATCGTCTTGCAATGAA ACAC ATACCGTTTTTC) and HLb(GCAAGATTTT TGCATCTTTAAAGGCCCCCCGATTTG) were used

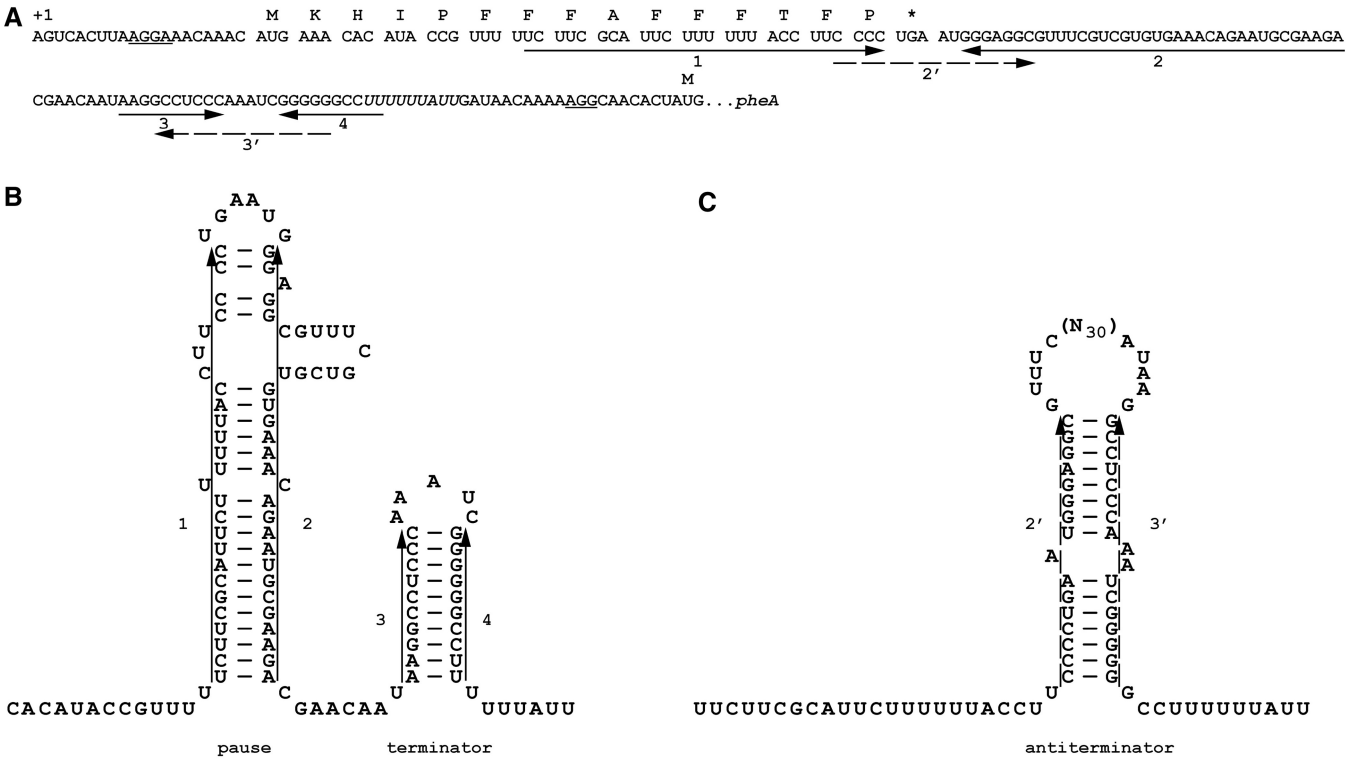

Figure 1. Regulatory mRNA structures in *pheA* expression. (**A**) mRNA sequence of the phenylalanine operon from the transcription initiation site to the *pheA* initiation codon. The coding sequence of *pheL* is separated into codons with the encoded amino acids indicated above. Shine–Dalgarno sequences 5′ of the *pheL* and *pheA* initiation codons are underlined. Solid arrows indicate nucleotides involved in formation of the pause and terminator RNA structures. Broken arrows indicate sequences that form the anti-termination secondary structure. The run of 'U's at which transcription attenuates is in italics. **(B)** RNA secondary structures promoting attenuation of transcription. Transcribing RNA polymerase proceeds rapidly to a site just after RNA segment 2 where it pauses (44) due to formation of the 1:2 RNA secondary structure. With ample phenylalanine, ribosomes do not pause at the phenylalanine codons but proceed to the stop codon. Once RNA polymerase continues transcription, the 3:4 attenuator generally forms and transcription terminates at the 3′-run of 'U's. **(C)** RNA secondary structure favouring transcription and synthesis of *pheA*. When insufficient phenylalanine leads to aminoacyl-tRNA^{Phe} becoming limiting and ribosome pausing at the run of 7 phenylalanine codons, RNA segment 2 is free to base pair with segment 3 which is transcribed once the polymerase is released from the pause. Formation of this antiterminator stem loop (2:3) permits transcription of the *pheA* coding sequence.

for the HLLH construct; HLa and LH2(GCATCTTTTC AGGGGAGG) were used for HLH5; HL1(CCTTCCCC TGAAAAGATGC) and HLb were used for HLH3. The 5′ *pdxH* sequence was amplified using the *pdxH* primer CT4t(ATAGGATCCAGCGTGAAAATGATGCGTGG AAG) and the *pheL*/*pdxH* chimera primers: LHb(CGGT ATGTGTTTCATTGCAAGACGATCAATC) for HLLH and HLH5 constructs and HL2(CTCCCATTCA GGGTGCAAG) for the HLH3 construct. The 3′ *pdxH* sequence was amplified using the *pdxH* primer CT4b(TAT GAATTCGAGTACCAGCGATTAAAGCAAG) and the *pheL*/*pdxH* chimera primers LHa(CGGGGGGCCT TTAAAGATGCAAAAATCTTGC) for HLLH and HLH3 and LH1(CCTTCCCCTGAAAAGATGC) for HLH5. DNA fragments from the first PCR step were gel-purified and a second PCR step was performed using CT4a and CT4b primers and the appropriate mix of fragments obtained in the first step. Changes were further introduced into the HLH5 construct to make MP, PM and constructs with separate codon mutations.

Construct CT5^{+} was made by cloning a synthetic oligonucleotide encompassing *pheL* sequence from the AUG codon to the CCC_UGAA stop, into the HindIII/ApaI sites between the GST-encoding sequence and *lacZ* of pSKAGS (48), so that *lacZ* was in the +1 frame relative to GST (frameshift reporter); for CT5IF, the cloned *pheL* sequence ended with CCC_TGG (to eliminate the UGA stop codon) and was placed in pSKAGS so that lacZ was in the 0-frame relative to the GST in-frame control).

For attenuation control experiments, the regulatory part of the phenylalanine operon was amplified by PCR from *E. coli* genomic DNA using the primers ATACTGC AGTTGACAGCGTGAAAACAGTAC and GGGATA TCGCATATGTGTCATAGTGTTGCCTTTTTGTTAT C. The PCR product was digested with NdeI and PstI restriction enzymes and cloned into corresponding sites of the pLA2 vector, removing the P_{araB} promoter (49). The resulting construct pPZWT has *lacZ* expression driven from the phenylalanine promoter and the initiation codon of *pheA*. Its expression is under *pheL* control and the regulatory secondary structures (Figure 7A).

Pulse-chase analysis in *E. coli*

Overnight cultures of DH5alpha *E. coli* strains expressing the appropriate construct were grown in MOPS-glucose (50) containing 100 μg/ml ampicillin and all amino acids

(150 μg/ml each) except methionine and tyrosine. They were diluted 1:50 in 300 μl of the same media. [Note, for the analysis shown in Figure 3C, phenylalanine was also omitted from the media used to generate material for the (−) lanes.] After 2-h incubation at 37°C, all cultures, except for the uninduced control, were induced with 2 mM IPTG for 10 min. The cells were pulse labelled for 2 min by addition of 7.5 μCi [^{35}S]-methionine in 30 μl media, chased for 2 min by addition of 30 μl cold methionine (50 mg/ml), chilled on ice and harvested by centrifugation. The pellets were resuspended in 50 μl Cracking Buffer (6 M urea, 1% sodium dodecyl sulphate, 50 mM Tris–HCl; pH 7.2) and heated at 95°C for 5 min. Aliquots of 5 μl were loaded on 4–12% NuPAGE Gels (Invitrogen Inc.) and electrophoresed, under conditions recommended by the manufacturer, in MOPS–SDS buffer (Invitrogen Inc.). Gels were exposed overnight and visualized with a Molecular Dynamics PhosphorImager. The amounts of termination and frameshift products were quantified by ImageQuant. The frameshifting efficiency was estimated as the ratio of the amount of frameshift product to the sum of the termination and frameshift products. At least three independent experiments were performed with each construct on separate days and the frameshifting levels presented in Figures 4–6 are the average values obtained, with the error bars indicating standard deviations.

Salmonella strains, genetic procedures and growth conditions

All strains (Table 1) are derivatives of *S. enterica* serovar Typhimurium strain LT2. As solid medium, LA (10 g of Trypticase peptone, 5 g of yeast extract, 5 g of NaCl and 15 g of agar per litre) was used. LB (51) and rich MOPS (50) were used as liquid media. Antibiotics were used at the following concentrations: carbenicillin (Cb): 50 mg/l; kanamycin (Km): 100 mg/l; chloramphenicol (Cm): 12.5 mg/l.

To transfer chromosomal markers or plasmids between strains of *S. enterica*, transductions were performed as described elsewhere (52) using a derivative of bacteriophage P22 containing the mutations HT105/I (53) and *int-201* (54).

proM2219 was isolated in a selection for frameshift suppressor derivatives of $tRNA^{Pro}_{cmo5UGG}$ that cause +1 frameshifting at CCC codons (38). It changes a G at position 31 in $tRNA^{Pro}_{cmo5UGG}$ to an A, resulting in an A–C mismatch instead of a G–C base pair between positions 31 and 39 in the lower part of the anticodon stem. *zhe-2533*::*cat* is a chloramphenicol-resistance cassette derived from plasmid pKD3 inserted 56 bp. downstream of the *proM* gene. It was used as selectable marker for strain constructions. Δ*proL* is a deletion of the gene encoding $tRNA^{Pro}_{GGG}$ (55), which also causes +1 frameshifting at CCC codons (56). The *pheR* mutant was constructed using λ-red recombineering (57). The *pheR* gene (one of the two *S. enterica* genes encoding $tRNA^{Phe}_{GAA}$) was replaced by the kanamycin-resistance cassette from plasmid pKD4, to yield *pheR*<>*kan* allele. To enable introduction of plasmid pPHWT (Km^R) into *pheR* strains, the kanamycin-resistance cassette was removed by expression of Flp recombinase from plasmid pCP20, resulting in the *pheR*<>*frt* allele (and the simultaneous conversion of *zhe-2533*::*cat* into *zhe-2533*::*frt*).

Table 1. Strains

Strain	Genotype	Reference or source
DH5alpha[a]	*endA1 recA1 relA1 gyrA96 hsdR17*($r_K^- m_K^+$) *phoA supE44 thi-1*Δ (*lacZYA-argF*)U169 Φ*80* Δ (*lacZ*)M15 F^-	Lab stock
GT8052	*attλ*::pPHWT *zhe-2533*::*cat*	This study
GT8053[b]	*attλ*::pPHWT ***proM2219***(G31A) *zhe-2533*::*cat*	This study
GT8054	*attλ*::pPHWT **Δ*proL*** *zhe-2533*::*cat*	This study
GT8055	*attλ*::pPHWT ***proM2219***(G31A) **Δ*proL*** *zhe-2533*::*cat*	This study
GT8056	*attλ*::pPHWT ***pheR***<>*frt* *zhe-2533*::*frt*	This study
GT8057	*attλ*::pPHWT ***pheR***<>*frt* ***proM2219***(G31A) *zhe-2533*::*frt*	This study
GT8058	*attλ*::pPHWT ***pheR***<>*frt* **Δ*proL*** *zhe-2533*::*frt*	This study
GT8059	*attλ*::pPHWT ***pheR***<>*frt* ***proM2219***(G31A) **Δ*proL*** *zhe-2533*::*frt*	This study
GT8060	CT5+/*zhe-2533*::*cat*	This study
GT8061	CT5+/***proM2219***(G31A) zhe-2533::cat	This study
GT8062	CT5+/**Δ*proL*** *zhe-2533*::*cat*	This study
GT8063	CT5+/***proM2219***(G31A) **Δ*proL*** *zhe-2533*::*cat*	This study
GT8064	CT5IF/*zhe-2533*::*cat*	This study
GT8065	CT5IF/***proM2219***(G31A) *zhe-2533*::*cat*	This study
GT8066	CT5IF/**Δ*proL*** *zhe-2533*::*cat*	This study
GT8067	CT5IF/***proM2219***(G31A) **Δ*proL*** *zhe-2533*::*cat*	This study
Plasmid	**Description**	
pPHWT	*pheL-pheA*::*lacZ* in pLA2, Km^R	This study
CT5+	*gst*::*pheL*::*lacZ* (*lacZ* in +1 frame compared to *gst*::*pheL*), Cb^R	This study
CT5IF	*gst*::*pheL*::*lacZ* (*lacZ* in 0 frame compared to *gst*::*pheL*), Cb^R	This study
pKD3	Template plasmid for chloramphenicol-resistance cassette	Datsenko & Wanner 2000
pKD4	Template plasmid for kanamycin-resistance cassette	Datsenko & Wanner 2000
pKD46	Helper plasmid for λ-red recombination	Datsenko & Wanner 2000
pINT-Ts	λ integrase helper plasmid	Hasan, Koob & Szybalski 1994
pCP20	Flp recombinase helper plasmid	Cherepanov & Wackernagel 1995

[a]DH5alpha is the only *E. coli* strain used in this study. All other strains listed in the table are *S. enterica* strains. [b]Mutations that are most experimentally relevant are in bold.

β-Galactosidase assays

Cultures for β-galactosidase assays were grown to mid-exponential phase ($OD_{600} \approx 0.4$–0.5 in a Beckman Coulter DU 730 spectrophotometer) in rich MOPS medium at 37°C and assayed (58), using the alternative protocol with SDS and chloroform instead of toluene to

permeabilize the cells. The frameshifting efficiencies reported in Figure 7A were calculated as below:

$$\text{FS}\ (\%) = 100 \times \frac{\text{beta gal.activity from CT5}^{+}}{\text{beta gal.activity from CT5IF}}$$

RESULTS

Conservation of *pheL* and the frameshift site

Since the *pheL* gene is very short, it is not annotated in the majority of sequenced genomes. To identify *pheL* genes, beyond those previously characterized in *E. coli*, *S. typhi*, *Klebsiella pneumoniae*, *Yersinia pestis*, *Erwinia carotovora*, *Serratia oneidensis* and *Vibrionales* (59), we extracted 200-nt upstream of annotated *pheA* genes in all sequenced bacteria. These were examined for the presence of at least five phenylalanine codons (where at least two of them are adjacent) occurring in-frame within a 60-nt region with no stop codons in the same translational phase. The intergenic regions between the resulting putative *pheL* genes and *pheA* genes were checked for the presence of a transcription termination signal and for potential termination/antitermination secondary structures. Those *pheL* genes that contain such signals and structures were classified as true positives. They were identified in a number of enteric bacteria including *Escherichia*, *Shigella*, *Salmonella*, *Erwinia*, *Klebsiella, Dickeya, Enterobacter, Citrobacter* and *Serratia* species (Figure 2). The *pheL* gene was also found in *Yersinia* species; however, the

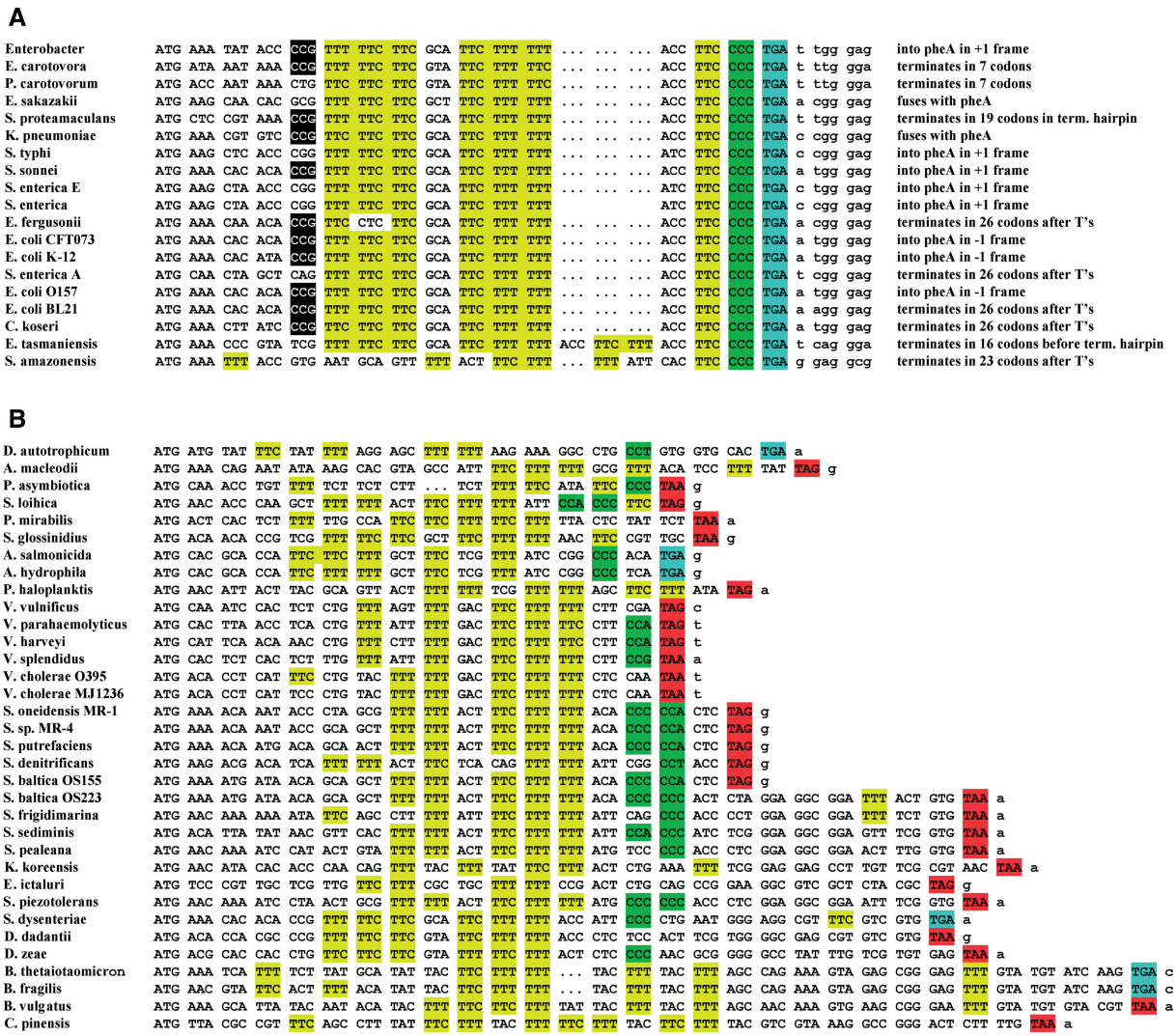

Figure 2. The *pheL* sequences from different bacteria. When the nucleotide sequence is identical in several strains or species (e.g. *E. coli* K-12 and *S. flexneri* 2a) it is included only once in the alignment. Phenylalanine codons are highlighted in yellow; TGA stop codons—blue; TAA and TAG stop codons—in red; and Proline codons—in green. The sequence sources are in the supplementary material. **(A)** Alignment of the *pheL* sequences terminating with CCC_UGA. The position of the −10 Proline codon that has a 2-fold effect on frameshifting, is indicated by black shading. For each sequence, continuation of translation in the +1 frame after CCC_UGA is described in terms of where it terminates relative to the frameshift site, termination hairpin or attenuation site (run of 'T's) or whether it proceeds past the *pheA* initiation codon and in which frame. **(B)** Alignment of the *pheL* sequences that do not end with CCC_UGA.

region between *pheL* and *pheA* is disrupted by an insertion sequence, IS1541, and the transcription attenuation mechanism might not be utilized to regulate *pheA* levels in *Yersinia*. In fact, in *Y. pestis*, the *pheA* gene itself is interrupted by another mobile element, IS100, and is unlikely to be functional. The *pheL* gene was also found in non-enteric bacteria, *Vibrio* and *Shewanella* species, of the gamma proteobacteria class. Only a few bacteria outside the proteobacteria class, *Bacteriodes* and *Desulfobacterium autotrophicum*, potentially use a *pheL* gene to regulate *pheA* expression.

All *pheL* genes can be separated into two groups. One group, limited to enterobacteria and *Shewanella amazonesis*, has highly conserved nucleotide sequence and all end with CCC_UGA (Figure 2A). In the other group, *pheL* sequences do not end with CCC_UGA and exhibit a lesser degree of conservation (Figure 2B). Plus-one frameshifting at CCC_UGA at the end of *pheL* would result in synthesis of a longer peptide. The amino acid sequence of this peptide, however, is not conserved among different enteric bacteria. The length of the peptide also differs. In *E. carotovora* [*Pectobacter carotovorum*, (60)] the stop codon in the +1 frame is only six codons 3′ of the shift site, on the 3′-side of the pause stem loop. In most other bacteria, the stop codon is located downstream of the pause structure (Figure 1B). In *Salmonella* species, the +1 frame stop codon overlaps with the AUG initiation codon of *pheA*. In *E. coli*, it is located 80-nt downstream of the *pheA* initiation codon. Moreover, in *K. pneumoniae* and *E. sakazakii*, +1 frameshifting on CCC_UGA would allow ribosomes to enter the *pheA* ORF. Since there are no stop codons separating the two genes in this frame, frameshifting in these bacteria would result in synthesis of a *pheL–pheA* fusion product. Thus, although the +1 frameshifting on CCC_UGA at the end of *pheL* is likely conserved in enteric bacteria, the length and the amino acid sequence of the synthesized peptide is not. Therefore, the peptide is unlikely to have a separate function.

Features in *pheL* that are responsible for frameshifting

The level of frameshifting in *pheL* gene decoding is much higher than in the 18 other *E. coli* genes that also end with CCC_UGA (1). An investigation of the features of *pheL* mRNA, and its encoded product responsible for the elevated frameshift levels, was undertaken. The frameshift reporter utilized encodes GST (glutathione-S-transferase)-MBP fusion protein (1). The *pheL* sequence to be tested was cloned so that its 5′ part is in-frame with GST, while MBP was placed in the +1 frame relative to GST. Thus, translation termination at CCC_UGA results in GST-PheL synthesis. If, however, +1 frameshifting at CCC_UGA occurs, then a GST–PheL–MBP fusion protein is synthesized. Frameshifting efficiencies were determined from expression levels of both proteins assayed by [^{35}S]-methionine pulse-chase labelling of total protein.

The possibility that sequence 3′ of CCC_UGA in *E. coli* *pheL* mRNA stimulates +1 frameshifting was addressed first. The original construct contained 33 nt of the natural sequence 3′ of CCC_UGA (1). Here (see Figure 3A), the 3′-sequence was shortened to only 9 nt 3′ of the CCC_UGA or extended to either 66 nt (including the RNA structures, but excluding the transcription termination site) or to 102 nt (including the start codon of *pheA*) 3′ of the CCC_UGA. No major change in frameshift efficiency was detected with the first two constructs (Figure 3B, first and second lane). With the construct that has 102 nt 3′ of CCC_UGA, the band corresponding to the frameshift product was faint (Figure 3B, third lane). Theoretically, sequences downstream of the CCC_UGA can preclude frameshifting. Alternatively, since the transcription termination site is located prior to the MBP sequence, lower levels of the full-length GST–pheL–MBP mRNA can be expected with this construct in the presence of phenylalanine (with high levels of the mRNA corresponding to GST-pheL). Consequently, a lower amount of the frameshift product with this construct can be explained not by lower frameshift levels *per se*, but by the absence of full-length mRNA containing sequences downstream of the frameshift site. To distinguish between the possibilities, the growth media was depleted of phenylalanine (to decrease transcription attenuation).

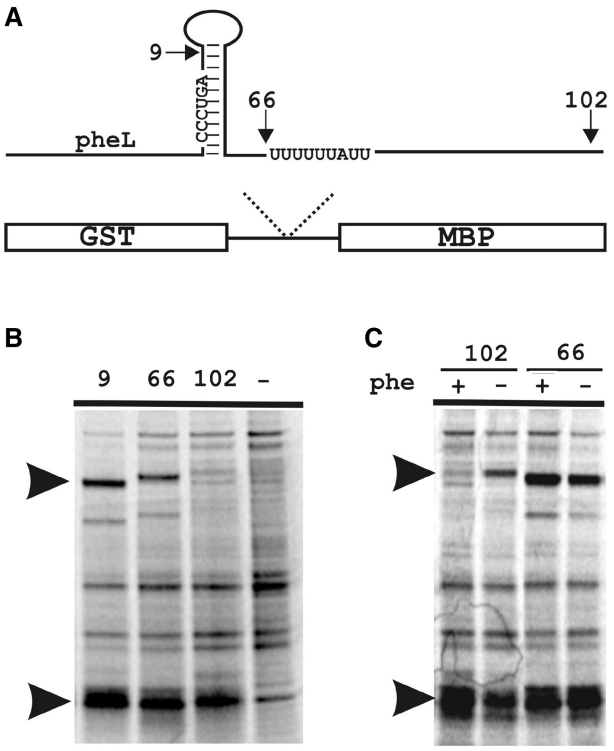

Figure 3. Influence of sequences 3′ of CCC_UGA frameshift site on efficiency of frameshifting in *pheL*. **(A)** Scheme of the reporter constructs. Arrows indicate the positions of the 3′-ends of sequences cloned in pGHM57 relative to regulatory structures and sequences in *pheL–pheA*. Numbers correspond to the number of nucleotides 3′ of CCC_UGA. **(B)** SDS–PAGE of protein products produced from *pheL* constructs with different 3′-lengths. The termination, GST-PheL, and the frameshift, GST–PheL–MBP, products are indicated by arrows. The (−) lane contains protein products from an uninduced control. **(C)** Effect of phenylalanine depletion on production of the frameshift product from the constructs with (102) and without (66) the transcription attenuation site.

This resulted in an increase of the frameshift product from the construct with the transcription termination site (Figure 3C, lanes 1 and 2). Phenylalanine depletion did not increase the amount of the frameshift product from the construct without the transcription termination site (Figure 3C, lanes 3 and 4). Thus, lower levels of GST–*PheL*–MBP frameshift product, from the construct containing the transcription attenuation site, are most likely due to a lower amount of the full-length mRNA produced. Therefore, most likely the 3′-context of the *pheL* mRNA does not influence, either positively or negatively, frameshifting on CCC_UGA.

To test this presumption, an additional experiment was performed. Sequence 5′ or 3′ of the *pheL* CCC_UGA was exchanged for the corresponding sequence 5′ or 3′ of CCC_UGA of the *pdxH* gene (Figure 4A). The *pdxH* coding sequence also ends with CCC_UGA, but a product derived from frameshifting during expression of this gene was not detected [(1); Figure 4B, lane 2]. First, the sequence from the *pheL* construct with 66 nt downstream of the frameshift site was moved in place of the CCC_UGA in the *pdxH*-containing construct (Figure 4A, third construct). This resulted in the 5′-*pdxH*–*pheL*–CCC_UGA–*pheL*–*pdxH*-3′ chimera construct (HLLH). The frameshifting efficiency on the CCC_UGA, with this context, was comparable with that exhibited by the original *pheL* construct (Figure 4B and C, lane 3). Next, either the 5′ or the 3′ *pheL* sequence was removed from the above chimera construct. The frameshifting efficiency with the construct 5′-*pdxH*–*pheL*–CCC_UGA–*pdxH*-3′ (HLH5; Figure 4B and C, lane 4) was comparable with that obtained for the wild-type *pheL*. The frameshifting efficiency, however, dropped to a marginal level with the construct 5′-*pdxH*–CCC_UGA–*pheL*–*pdxH*-3′ (HLH3; Figure 4B and C, lane 5). This confirms that the main stimulatory signals for frameshifting are located 5′ of the sequence CCC_UGA in *pheL*, with the identity of the 3′-sequence being unimportant.

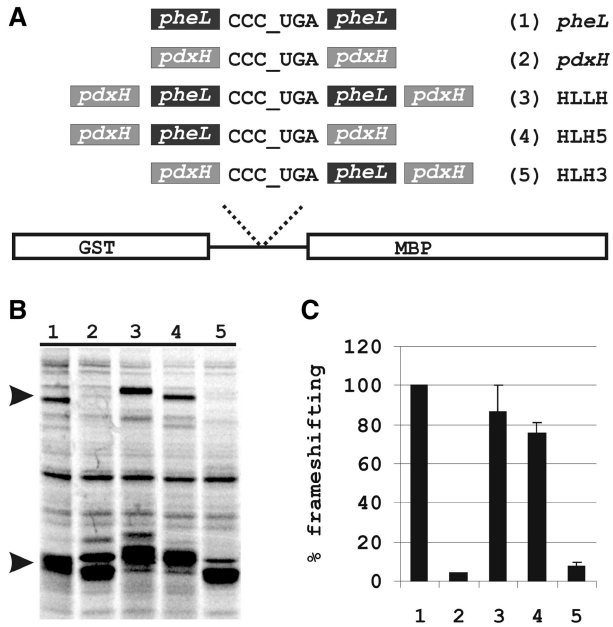

Figure 4. Effect of *pheL* and *pdxH* 5′ and 3′ sequences on frameshifting at CCC_UGA. **(A)** Scheme of the analysed constructs. **(B)** SDS–PAGE of proteins expressed from *pheL*, *pdxH*, HLLH, HLH5 and HLH3 constructs. The termination, GST–PheL, and the frameshift, GST–PheL–MBP, products are indicated by arrows. **(C)** Quantitation of frameshifting efficiency during translation of *pheL*, *pdxH*, HLLH, HLH5 and HLH3 constructs from (B). The level of frameshifting obtained with the WT *pheL* sequence is set to 100% of wild-type.

The 5′-stimulator can act at the nucleotide level and/or at the level of the encoded amino acids in the nascent peptide. To distinguish between these possibilities, two types of experiments were performed with the HLH5 construct described above. In one set of experiments, single nucleotide insertions and deletions were introduced in the *pheL* sequence so that ribosomes translating the 5′-region were routed in and out of different frames. The synthesized peptide thus had a completely different amino acid sequence, but the nucleotide sequence was maximally preserved. Two series of constructs were made for this purpose: PM0-3 and MP0-3 (Figure 5A). In the PM series, a nucleotide was inserted at the start codon of *pheL* and a compensatory nucleotide deletion was made prior to the sequence corresponding to the CCC_UGA frameshift site. In the MP series, 4 nt, ATGA, were deleted at the beginning of the *pheL* and a compensatory insertion prior to the frameshift site routed ribosomes back into the zero frame. Numbers 0 through 3 indicate the number of wild-type *pheL* codons in a construct immediately 5′ of the CCC_UGA. Constructs PM0 and MP0, in which ribosomes were routed back in-frame just prior to CCC_UGA, showed a significant drop in frameshifting efficiency, to 10 and 30% of WT *pheL*, respectively (Figure 5B and C). The frameshifting efficiency gradually increased with the addition of WT codons upstream of the CCC_UGA. Constructs MP3 and PM3, which had three wild-type codons upstream of the CCC_UGA, exhibited wild-type levels of frameshifting.

In the other set of mutants, codons throughout the *pheL* ORF in the HLH5 reporter were altered separately, or in blocks, to either synonymous or non-synonymous codons (Figure 6A). None of the changes to synonymous codons affected frameshifting efficiency (Figure 6B–G), while non-synonymous changes of a few codons had a much stronger effect. The identity of the amino acid, phenylalanine, encoded 5′ adjacent to CCC_UGA is crucial for the observed frameshifting levels (Figure 6D and E). Only changing the third position of this phenylalanine codon, UUC, to the synonymous UUU did not alter the frameshifting level, whereas changing it to UUG (leucine) reduced frameshifting to <20% of the WT level (as did changing the UUC to CGC, arginine). Changing just the second nucleotide of UUC so that the codon became UAC (tyrosine) reduced the frameshifting level to ~50% of WT (a similar level to changing the UUC to a different tyrosine codon (UAU)). Changing just the first nucleotide of UUC so the codon became CUC (leucine) also drastically reduced frameshifting. Thus, frameshifting was unaffected by minimal changes that preserved codon meaning but reduced when the identity of the encoded amino acid was altered. Changing the ACC (threonine) codon at the −2 position with respect to the CCC_UGA

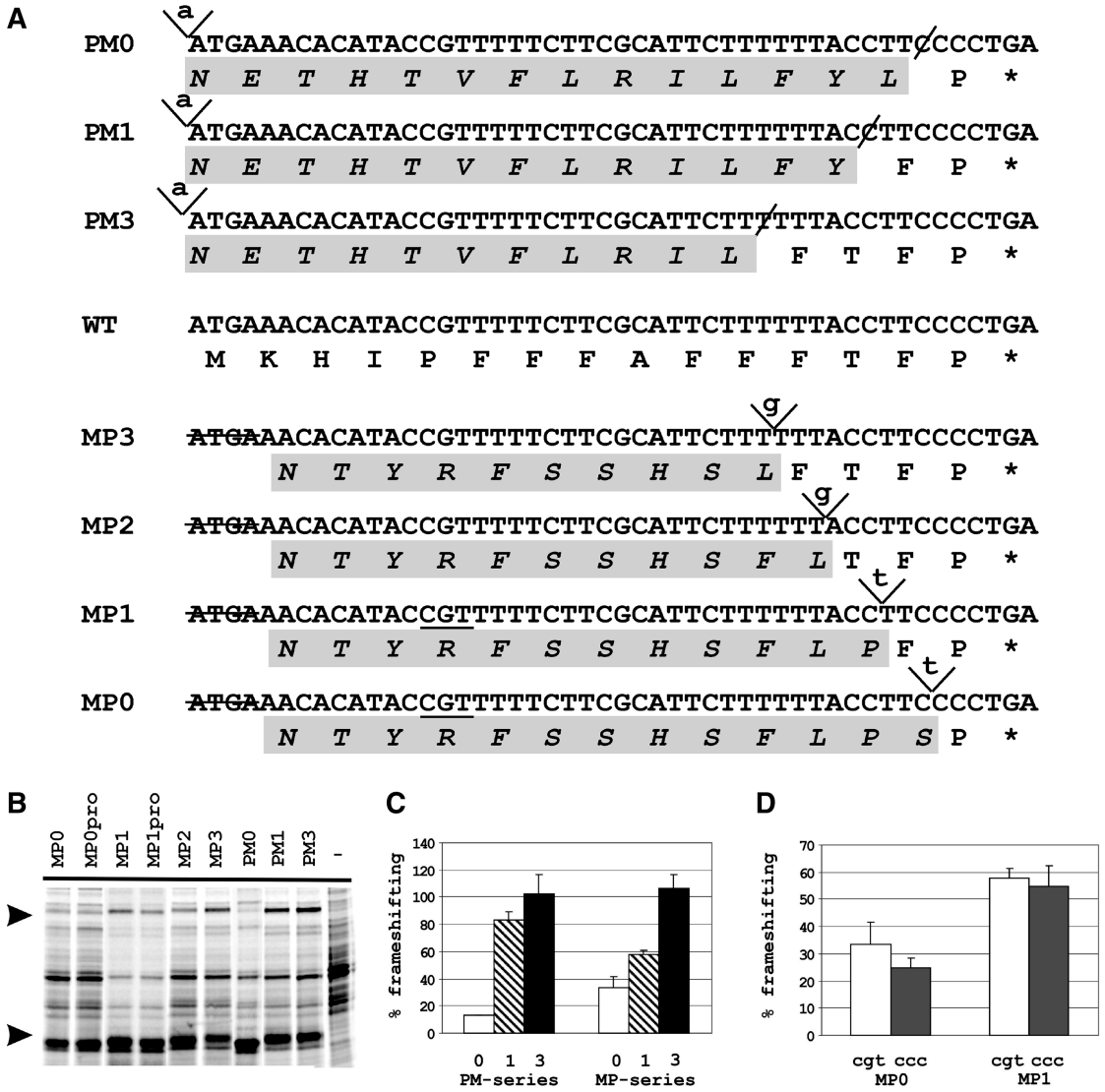

Figure 5. Analysis of the 5′-stimulator in *pheL* by insertion/deletion experiment. **(A)**. Sequences of MP and PM constructs. In the MP and PM series inserted nucleotides are italicized and underlined, deleted nucleotides are designated by a hyphen (-). Sequence in capital letters represents the preserved codons of the wild-type *pheL*; lower case letters are used for the 'out-of-frame' sequence altered at the level of the encoded amino acids. The codon at position −10 in the MP0 and MP1 constructs is underlined. **(B)** SDS–PAGE of the protein products expressed from MP, MPpro (see text) and PM constructs. The termination, GST–PheL, and the frameshift, GST–PheL–MBP, products are indicated by arrows. The (−) lane contains protein products from an uninduced control. **(C)** Quantitation of frameshifting levels obtained with the MP and PM constructs relative to that of WT *pheL*, which is set at 100%. The frameshifting efficiency with the MP2 construct was about the same as with MP1 and is not shown separately. **(D)** Comparison of frameshifting efficiencies in the MP0 and MP1 (CGT) constructs with MP0pro and MP1pro (CCC).

to the synonymous codon ACA did not affect frameshifting levels (Figure 6B and C), but the non-synonymous alteration to a glycine (GGC) codon resulted in an ~50% increase in frameshifting (Figure 6B and C).

A single nucleotide change in the 10th codon 5′ of the shift site, CCG to CCU (both proline) did not affect the frameshifting level, whereas changing the CCG to ACG (threonine) or CUG (serine) or any of the other four non-proline codons tested, led to a 2-fold diminution of frameshifting. The other changes that resulted in alteration of the encoded amino acid also resulted in a similar reduction of frameshifting (Figure 6F and G). [The role of the proline codon at position −10 was investigated further by changing it to a different proline codon, CCC, in the MP0 and MP1 constructs, that encode a scrambled peptide sequence (see above and Figure 5A). No significant change in frameshifting efficiency was observed (Figure 5B and D).]

Taken together, these results provide suggestive evidence that the frameshift stimulator in the *pheL* gene is mainly composed of the amino acids encoded in the mRNA sequence. A different peptide stimulator acts in the 50% efficient translational bypassing of the 50 non-coding nucleotides between codons 46 and 47 of bacteriophage T4 gene *60* mRNA (61–63). The nascent peptide stimulator acts to promote dissociation of peptidyl-tRNA anticodon pairing to codon 46 (4,64). Although in bypassing the anticodon re-pairing to

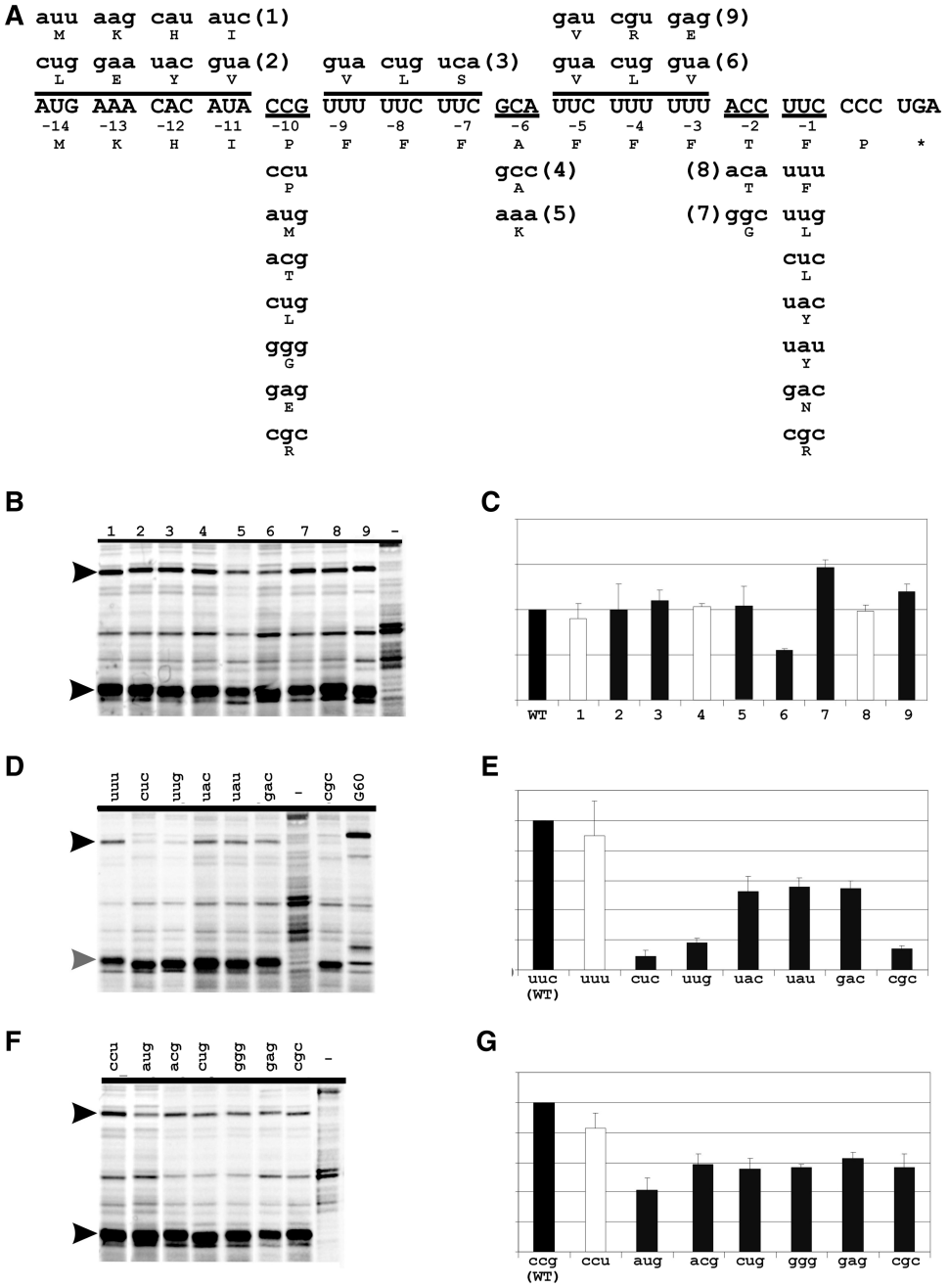

Figure 6. Analysis of the peptide stimulator by synonymous and non-synonymous mutations in *pheL*. In (**C**), (**E**) and (**G**), white bars correspond to synonymous and grey to non-synonymous mutations. In (**B**), (**D**) and (**F**) the (−) lane contains protein products from an uninduced control. The termination, GST–PheL, and the frameshift, GST–PheL–MBP, products are indicated by arrows. **(A)** Peptide mutations (in lower case letters). The wild-type *pheL* sequence is in capital letters. Negative numbers above the codons of the *pheL* peptide represent the position of a codon relative to the CCC_UGA. Codons mutated simultaneously in a single construct are under- (or over-) lined. Nucleotides differing from the wild-type sequence are underlined. Numbers in brackets indicate the number of a construct and correspond to the numbers of constructs analysed in (B) and (C). (B) SDS–PAGE of protein products from pulse-chase experiment with the 'numbered' constructs from (A). (C) Relative quantitation of frameshifting efficiencies in the constructs analysed in (B). All frameshifting efficiencies are relative to that of WT *pheL*, which is set at 100%. (D) SDS–PAGE of proteins from pulse-chase analysis of constructs with changes in the −1 codon of *pheL*. Grey arrow indicates a putative tmRNA-tagged termination product discussed in the text. (E) Relative quantitation of frameshifting efficiencies in the constructs analysed in (D). (F) SDS–PAGE of proteins produced from the constructs with changes in the –10 codon of *pheL*. **(G)** Relative quantitation of frameshifting efficiencies in the constructs analysed in (F).

mRNA is not at an overlapping codon, and in gene *60* bypassing the sequence of the coding gap is not scanned, the key feature of P-site codon:anticodon dissociation is shared by at least most cases of programmed frameshifting. Accordingly, we tested whether the gene *60* nascent peptide would stimulate +1 frameshifting at CCC_UGA. The 5′ *pheL* sequence in the HLH5 construct was substituted with the 135 nt (45 codons) preceding the

codon 46 GGA take-off site of gene *60*. Remarkably, ~60% +1 frameshifting was observed on CCC_UGA with the gene *60* nascent peptide stimulator (Figure 6D, right-most lane).

Role of frameshifting at the end of *pheL*

The amino acid sequence encoded by the new frame after the +1 ribosomal frameshift is not conserved (Figure 2A). Sequence comparisons, therefore, do not provide support for a functional role for the protein product. Frameshifting at the end of *pheL* might influence the levels of *pheA* expression. Ribosomes that escape the *pheL* stop codon and continue translating the *pheL–pheA* mRNA, would preclude formation of the attenuator structure and might contribute to antitermination of *pheA* transcription.

To test this, we availed of mutants that showed higher or lower frameshifting than wild-type at CCC_UGA. In *S. enterica*, +1 frameshifting at CCC codons is known to be enhanced by various mutations affecting $tRNA^{Pro}_{GGG}$ (*sufB2*, Δ*proL*) or $tRNA^{Pro}_{GGG}$(*sufA6*). Most, or all, of the frameshifting in these mutants occurs when the structurally normal near-cognate peptidyl-tRNA, $tRNA^{Pro}_{cmo5UGG}$, occupies the P-site of the ribosome with a CCC codon (39,56,65). In addition, a number of mutants of $tRNA^{Pro}_{cmo5UGG}$ (*proM*) have recently been isolated as +1 frameshift mutant suppressors acting at CCC codons (38).

We used plasmids CT5+ and CT5IF, which carry *gst–pheL–lacZ* fusions, to determine the frequencies of +1 frameshifting in the various mutants. In these plasmids, *lacZ* is in the +1 frame (CT5+), or the 0 frame (CT5IF), compared with *gst–pheL*, and frameshifting was monitored by β-galactosidase assays. The Δ*proL* mutants caused a 4.4-fold increase in frameshifting compared with wild-type (1.9% in wt, versus 8.4% in Δ*proL*), and the most efficient *proM* mutant, *proM2219*, increased frameshifting 2.5-fold to 4.7% (Figure 7A). The combination of *proM2219* and Δ*proL* led to a 14.5-fold increase in frameshifting (27.5%) as the mutant $tRNA^{Pro}_{cmo5UGG}$ did not have to compete with any other tRNA for CCC-codons. Other frameshift suppressor mutants (*sufA6*, *sufB2*) were tested and caused frameshifting at about the same level as the Δ*proL* mutant. A *cmoB2* mutant, known to decrease frameshifting at other CCC_UGA sites, did not decrease frameshifting below the wild-type level (data not shown).

To test the functional effect of frameshifting at the *pheL* CCC_UGA on attenuation control, a new reporter was constructed. The nucleotide sequence encompassing the promoter region, *pheL* and the *pheL–pheA* intergenic region, including the AUG start codon of *pheA*, was cloned upstream of *lacZ* in the pLA2 vector (49). The *lacZ* gene was placed in-frame with the *pheA* initiation AUG. Therefore, in this construct, pPZWT, β-galactosidase expression was driven by the phenylalanine operon promoter and was under similar control (such as *pheL* regulatory RNA structures and an initiation stimulatory Shine–Dalgarno) as endogenous *pheA*.

Whether wild-type cells are grown on minimal media or minimal media supplemented with the aromatic amino acids, the levels of *pheA* expression remain essentially the same (41). The *pheA* gene can be 2- to 20-fold de-repressed by phenylalanine starvation in chemostat conditions and/or with specific mutants (66,67), such as those that cause decreased amount or function of Phe-$tRNA^{Phe}$. One such mutant is *miaA*, which lacks the $ms^2i(o)^6A37$ modification present in $tRNA^{Phe}$ (68). Another is *pheR*, which was first thought to be a repressor protein, but was later shown to lack one of the two genes encoding $tRNA^{Phe}$ (69). In our hands, the *miaA1* mutation did not cause any de-repression of β-galactosidase activity from the pPHWT plasmid (data not shown), whereas with

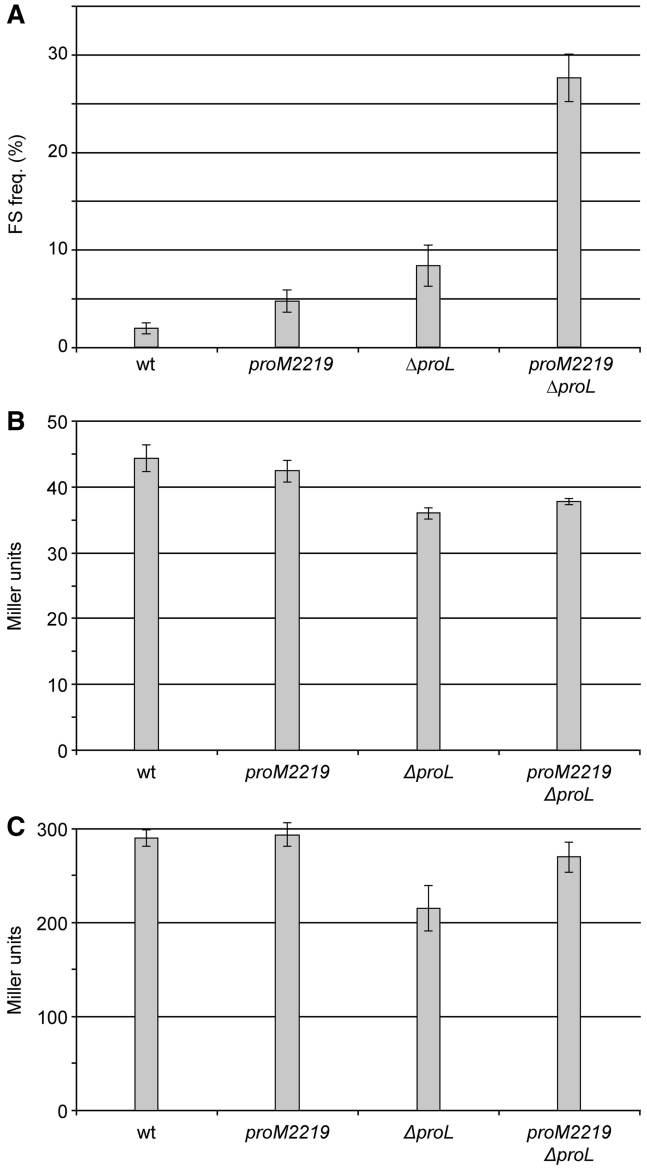

Figure 7. Frameshifting in *pheL* decoding in *Salmonella* mutants and effect on attenuation of transcription. **(A)** Efficiency of frameshifting during *pheL* decoding in different mutants. The values are averages of two experiments with at least three independent cultures of each. **(B)** β-Galactosidase activities of the *pheA–lacZ* fusion (in pPHWT) during repressed conditions (wild-type *pheR*); **(C)** de-repressed conditions (*pheR<>frt* mutant). The values are averages of four independent cultures with error bars representing standard deviation from the mean.

a *pheR* mutant, activity was de-repressed 6- to 7-fold (Figure 7B and C). None of the tested frameshift suppressors (*proM2219* and Δ*proL* single mutants or *proM2219* Δ*proL* double mutants) caused any significant change in β-galactosidase activity either during repressing conditions (*pheR*$^+$; Figure 7B) or de-repressing conditions (*pheR<>frt*; Figure 7C). From these data, we conclude that frameshifting at the end of *S. enterica pheL* does not contribute significantly to the regulation of *pheA*, and is, therefore, not likely to be a true case of utilized frameshifting.

DISCUSSION

pheL sequence conservation

Nearly half of the identified *pheL* coding sequences terminate with UGA, and in these there is an impressive, although not absolute correlation, with both phenylalanine and proline codons being encoded by the penultimate and 5′-adjacent codons to this stop codon. This contrasts with those that terminate with UAG or UAA. The 19 sequences in Figure 2A are highly related while the majority of those in Figure 2B are quite diverse. Some role for poor termination in those depicted in Figure 2A likely awaits discovery. It is tempting to deduce that selection has not only been acting at the amino acid level. In all 19 *pheL* genes in Figure 2A, the nucleotide sequence is UUC_CCC_UGA whereas in all other positions, except for the initiator AUG and regulatory phenylalanine codons, variations do occur. Our results obtained with *S. enterica* mutants, however, argue against selection of this sequence for frameshifting purposes.

While the length of *pheL* and identity of most of its codons are variable, as expected, the regulatory phenylalanine codons are substantially conserved (Figure 2). A proline codon commonly occurs in the 15th (+ or −1) position and may, via a pausing effect, also be important for the regulation. However, this region is key for formation of both antitermination and pause-structure hairpins and the 'C's of the proline codons are complementary to the downstream 'G's. In the few sequences that do not have proline codons in those positions, the nucleotide sequence is pyrimidine rich, probably to allow the base pairing with the 'G's. Changes to purine nucleotides in this region are highly unlikely, because they would require compensatory changes on the other sides of both antitermination and pause stems and those requiring further compensatory changes in the terminator hairpin.

pheL and *tnaC*

Despite the obvious differences between the biosynthesis of phenylalanine and catabolism of tryptophan, some comparison between the present *pheL* work and *tnaC* is merited. Although C-terminal proline is important for *E. coli tnaC* expression, it is also just semi-conserved (18). In *Proteus vulgaris*, there are two extra codons 3′ of the proline codon before the termination codon and, yet, pausing still occurs there. Among the 34 sequences listed in Figure 2B, 20 have a proline codon close to the same position in relation to the initiation codon as do the 19 in Figure 2A that terminate with UUC_CCC_UGA, although several have extra codons after the proline codon. Of those, seven unique but related sequences from the *Shewanella* species, have tandem proline codons. Nevertheless, there are no tandem proline codons in any of the 23 sequences when a proline codon is 5′-adjacent to a stop codon. It is tempting to infer similarities to the pausing location in *Proteus vulgaris tnaC*. [An Asp, eight residues N-terminal of the proline residue, is important for *tnaC* pausing (18). There is no Asp codon in *pheL* and, in the great majority of the sequences, there is a Phe codon at the equivalent position to the Asp codon in *tnaC,* or else adjacent to it.] *E. coli tnaC* terminates with CCU_UGA rather than CCC_UGA in *E. coli pheL* (the 3′-proline codon in 9 of the 34 sequences in Figure 2B is not CCC by contrast to the likely frameshift-prone CCC-containing sequences in Figure 2A). The tRNA that decodes CCC also decodes CCU and a possible role for frameshifting in expression of *E. coli tnaC*, was considered at an early stage in the analysis of that operon, but dismissed (70). In the *pheL* context when CCC_UGA is mutated to CCU_UGA, the frameshifting level dropped by two-thirds (1). More importantly, the N-terminal amino acid adjacent to Pro in TnaC is Arg, and placing an arginine codon at the corresponding position of *pheL* virtually abolished frameshifting. Thus, the context of the termination codon in TnaC makes frameshifting in this gene highly unlikely. More than a decade of elegant work has revealed the nature of *tnaC* expression. By contrast, many aspects of *pheL* expression, such as possible mRNA cleavage, remain to be studied and are outside the scope of the present work.

Non-standard translational events other than frameshifting

Inefficient translational termination context in the *pheL* gene can result not only in frameshifting but other non-standard translational events. Stop-codon readthrough, measured by pulse-chase analysis with the GST–*pheL*–MBP construct, occurs at an ∼4% level in *E. coli* (data not shown). The GST–PheL termination product has different mobility with different constructs (i.e. in Figures 5 and 6). For example, in Figure 6D the first lane has products from the HLH5 construct in which the UUC codon preceding CCC_UGA was changed to another phenylalanine codon, UUU. The band corresponding to the termination product migrates slower than in the next lane, which contains products from the construct in which the UUC codon was changed to a CUC leucine codon. The likely explanation for this phenomenon is that phenylalanine in the penultimate position, unlike leucine, hinders termination and promotes alternative translational events such as frameshifting and tmRNA tagging (71). A tmRNA-tagged termination product would be 10 amino acids longer and would migrate slightly slower than the untagged version. The duration of the pulse-chase experiment is only 1–2 min and at least part of any tagged product may not be degraded. Our attempts to purify a GST-PheL tmRNA-tagged product from either wild-type cells or tmRNA$^-$ cells expressing tmRNA encoding a 6-histidine

tag (SsrA-H_6), (72), failed. Nevertheless, a 6His-tagged product was detected by western blotting when GST-*pheL* fusion was expressed in the SsrA-H_6 cells (O.G. and N.M. Wills, unpublished data). The tagging efficiency was comparable with the control GST-*ybeL* fusion; The YbeL protein also has a proline at C-terminus and was first identified as a target for tmRNA-tagging (17,72). Thus, the CCC_UGA sequence at the end of *E. coli pheL* is a site for readthrough and tmRNA tagging as well as frameshifting.

Frameshifting in *pheL* decoding

Frameshifting at the CCC_ UGA in decoding *E. coli* K12 *pheL* occurs with 15% efficiency, whereas in wild-type *S. enterica* the comparable frameshifting efficiency is much lower, ~1.9% (Figure 7A). The reason for the difference between the *pheL* frameshifting levels in *S. enterica* and *E. coli* has not been studied, but two aspects merit consideration. *E. coli* is 15% more negatively supercoiled than *S. enterica* with significant effects on gene expression (73). Increasing the level of negative supercoiling of the *S. enterica* promoter of the 4-gene tRNA operon that includes *proM*, the gene for the near-cognate proline $tRNA_{cmo5UGG}$ (74), leads to elevated levels of this tRNA (75). Perhaps there is an elevated ratio of *proM*/*proL*-encoded tRNAs in *E. coli* compared with *S. enterica* and, if so, it would lead to more frameshifting; however, it is unknown whether the relative expression of *proM*/*proL* has been adapted to compensate for the different levels of supercoiling in different organisms. Another, even more relevant, consideration is a key difference between the release factors 2 of *S. enterica* and *E. coli K12* that results in reduced termination efficiency of the latter (76,77). This is correlated with enhanced frameshifting at a codon 5′-adjacent to UGA (78).

Björnsson *et al.* (15) showed that phenylalanine in the penultimate position diminishes termination efficiency and suggested that hydrophobic and acidic amino acids, in general, have this effect. Indeed, when the *pheL* penultimate sense codon, that for phenylalanine, was substituted with a tyrosine (hydrophobic) or aspartic acid (acidic) codon, the level of frameshifting was reduced by half, which was still much higher than the undetectable level when an arginine (basic) codon is present (Figure 6D and E). Nevertheless, changing the phenylalanine codon preceding CCC_UGA in *pheL* to a tyrosine codon did result in a 2-fold reduction of frameshifting efficiency (Figure 6D and E), suggesting that the antitermination effect of the penultimate amino acid is not solely dependent on its hydrophobic properties.

Programmed autoregulatory frameshifting is utilized in *E. coli* release factor 2 mRNA decoding. The sequences features involved in this frameshifting incidentally cause some level of internal initiation, but the process has no functional implications (79). By contrast, the present analysis suggests that in *pheL* decoding it is the frameshifting that is incidental to sequence features selected to influence attenuation via their effect on termination. The general reason for the extent of the nucleotide conservation 5′ of *pheL* genes that terminate with UGA remains for future work to discern.

Searches for new cases where frameshifting is utilized for gene expression (80) should be (i) mindful that some programmed and high-level frameshifting may be incidental and (ii) that nascent peptide stimulators may be awaiting discovery even if the levels of frameshifting attained does not approach the 60% found here with a heterologous nascent peptide stimulator.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

We are grateful to Norma Wills for her assistance and suggestions on experimental procedures, to Dr James Pittard for his crucial advice on de-repression conditions for phenylalanine operon and for providing the AB1360 strain and to the *E. coli* Genetic Stock Center for providing pLA2 and pINT-ts vectors.

FUNDING

Swedish Science Council and the Carl Trygger Foundation (to G.R.B.); National Institute of Health (grant number GM079523 to J.F.A.); Science Foundation Ireland (to P.V.B. and J.F.A.). Funding for open access charge: Science Foundation Ireland (to J.F.A.).

Conflict of interest statement. None declared.

REFERENCES

1. Gurvich,O.L., Baranov,P.V., Zhou,J., Hammer,A.W., Gesteland,R.F. and Atkins,J.F. (2003) Sequences that direct significant levels of frameshifting are frequent in coding regions of *Escherichia coli*. *EMBO J.*, **22**, 5941–5950.
2. Gurvich,O.L., Baranov,P.V., Gesteland,R.F. and Atkins,J.F. (2005) Expression levels influence ribosomal frameshifting at the tandem rare arginine codons AGG_AGG and AGA_AGA in *Escherichia coli*. *J. Bacteriol.*, **187**, 4023–4032.
3. Ivanov,I.P. and Atkins,J.F. (2007) Ribosomal frameshifting in decoding antizyme mRNAs from yeast and protists to humans: close to 300 cases reveal remarkable diversity despite underlying conservation. *Nucleic Acids Res.*, **35**, 1842–1858.
4. Herr,A.J., Nelson,C.C., Wills,N.M., Gesteland,R.F. and Atkins,J.F. (2001) Analysis of the roles of tRNA structure, ribosomal protein L9, and the bacteriophage T4 gene *60* bypassing signals during ribosome slippage on mRNA. *J. Mol. Biol.*, **309**, 1029–1048.
5. Vazquez-Laslop,N., Thum,C. and Mankin,A.S. (2008) Molecular mechanism of drug-dependent ribosome stalling. *Mol. Cell*, **30**, 190–202.
6. Oliver,D., Norman,J. and Sarker,S. (1998) Regulation of *Escherichia coli secA* by cellular protein secretion proficiency requires an intact gene X signal sequence and an active translocon. *J. Bacteriol.*, **180**, 5240–5242.
7. Gong,F. and Yanofsky,C. (2002) Instruction of translating ribosome by nascent peptide. *Science*, **297**, 1864–1867.
8. Cruz-Vera,L.R., Yang,R. and Yanofsky,C. (2009) Tryptophan inhibits *Proteus vulgaris* TnaC leader peptide elongation, activating tna operon expression. *J. Bacteriol.*, **191**, 7001–7006.

9. Nakatogawa,H. and Ito,K. (2002) The ribosomal exit tunnel functions as a discriminating gate. *Cell*, **108**, 629–636.
10. Yap,M.N. and Bernstein,H.D. (2009) The plasticity of a translation arrest motif yields insights into nascent polypeptide recognition inside the ribosome tunnel. *Mol. Cell*, **34**, 201–211.
11. Tanner,D.R., Cariello,D.A., Woolstenhulme,C.J., Broadbent,M.A. and Buskirk,A.R. (2009) Genetic identification of nascent peptides that induce ribosome stalling. *J. Biol. Chem.*, **284**, 34809–34818.
12. Seidelt,B., Innis,C.A., Wilson,D.N., Gartmann,M., Armache,J.P., Villa,E., Trabuco,L.G., Becker,T., Mielke,T., Schulten,K. *et al.* (2009) Structural insight into nascent polypeptide chain-mediated translational stalling. *Science*, **326**, 1412–1415.
13. Cridge,A.G., Major,L.L., Mahagaonkar,A.A., Poole,E.S., Isaksson,L.A. and Tate,W.P. (2006) Comparison of characteristics and function of translation termination signals between and within prokaryotic and eukaryotic organisms. *Nucleic Acids Res.*, **34**, 1959–1973.
14. Janssen,B.D. and Hayes,C.S. (2009) Kinetics of paused ribosome recycling in *Escherichia coli*. *J. Mol. Biol.*, **394**, 251–267.
15. Björnsson,A., Mottagui-Tabar,S. and Isaksson,L.A. (1996) Structure of the C-terminal end of the nascent peptide influences translation termination. *EMBO J.*, **15**, 1696–1704.
16. Cao,J. and Geballe,A.P. (1996) Inhibition of nascent-peptide release at translation termination. *Mol. Cell Biol.*, **16**, 7109–7114.
17. Hayes,C.S., Bose,B. and Sauer,R.T. (2002) Proline residues at the C terminus of nascent chains induce SsrA tagging during translation termination. *J. Biol. Chem.*, **277**, 33825–33832.
18. Cruz-Vera,L.R. and Yanofsky,C. (2008) Conserved residues Asp16 and Pro24 of TnaC-tRNAPro participate in tryptophan induction of *tna* operon expression. *J. Bacteriol.*, **190**, 4791–4797.
19. Weiss,R.B., Dunn,D.M., Atkins,J.F. and Gesteland,R.F. (1987) Slippery runs, shifty stops, backward steps, and forward hops: −2, −1, +1, +2, +5, and +6 ribosomal frameshifting. *Cold Spring Harb. Symp. Quant. Biol.*, **52**, 687–693.
20. Weiss,R.B., Dunn,D.M., Atkins,J.F. and Gesteland,R.F. (1990) Ribosomal frameshifting from −2 to +50 nucleotides. *Prog. Nucleic Acid Res. Mol. Biol.*, **39**, 159–183.
21. de Smit,M.H., van Duin,J., van Knippenberg,P.H. and van Eijk,H.G. (1994) CCC.UGA: a new site of ribosomal frameshifting in *Escherichia coli*. *Gene*, **143**, 43–47.
22. Vilbois,F., Caspers,P., da Prada,M., Lang,G., Karrer,C., Lahm,H.W. and Cesura,A.M. (1994) Mass spectrometric analysis of human soluble catechol O-methyltransferase expressed in *Escherichia coli*. Identification of a product of ribosomal frameshifting and of reactive cysteines involved in S-adenosyl-L-methionine binding. *Eur. J. Biochem.*, **222**, 377–386.
23. O'Connor,M. (2002) Imbalance of tRNA(Pro) isoacceptors induces +1 frameshifting at near-cognate codons. *Nucleic Acids Res.*, **30**, 759–765.
24. Ivanov,I.P., Gesteland,R.F. and Atkins,J.F. (2000) Antizyme expression: a subversion of triplet decoding, which is remarkably conserved by evolution, is a sensor for an autoregulatory circuit. *Nucleic Acids Res.*, **28**, 3185–3196.
25. Dorscht,J., Klumpp,J., Bielmann,R., Schmelcher,M., Born,Y., Zimmer,M., Calendar,R. and Loessner,M.J. (2009) Comparative genome analysis of *Listeria* bacteriophages reveals extensive mosaicism, programmed translational frameshifting, and a novel prophage insertion site. *J. Bacteriol.*, **191**, 7206–7215.
26. Zimmer,M., Sattelberger,E., Inman,R.B., Calendar,R. and Loessner,M.J. (2003) Genome and proteome of *Listeria monocytogenes* phage PSA: an unusual case for programmed +1 translational frameshifting in structural protein synthesis. *Mol. Microbiol.*, **50**, 303–317.
27. Fortier,L.C., Bransi,A. and Moineau,S. (2006) Genome sequence and global gene expression of Q54, a new phage species linking the 936 and c2 phage species of *Lactococcus lactis*. *J. Bacteriol.*, **188**, 6101–6114.
28. Auzat,I., Droge,A., Weise,F., Lurz,R. and Tavares,P. (2008) Origin and function of the two major tail proteins of bacteriophage SPP1. *Mol. Microbiol.*, **70**, 557–569.
29. Liao,P.Y., Choi,Y.S. and Lee,K.H. (2009) FSscan: a mechanism-based program to identify +1 ribosomal frameshift hotspots. *Nucleic Acids Res.*, **37**, 7302–7311.
30. Weiss,R.B., Dunn,D.M., Shuh,M., Atkins,J.F. and Gesteland,R.F. (1989) *E. coli* ribosomes re-phase on retroviral frameshift signals at rates ranging from 2 to 50 percent. *New Biol.*, **1**, 159–169.
31. Fayet,O. and Prère,M.-F. (2010) In Atkins,J.F. and Gesteland,R.F. (eds), *Recoding: Expansion of Decoding Rules Enriches Gene Expression*. Springer, New York and Heidelberg, pp. 259–281.
32. Tsuchihashi,Z. and Brown,P.O. (1992) Sequence requirements for efficient translational frameshifting in the *Escherichia coli dnaX* gene and the role of an unstable interaction between tRNA(Lys) and an AAG lysine codon. *Genes Dev.*, **6**, 511–519.
33. Murphy,F.V., Ramakrishnan,V., Malkiewicz,A. and Agris,P.F. (2004) The role of modifications in codon discrimination by tRNA(Lys)UUU. *Nat. Struct. Mol. Biol.*, **11**, 1186–1191.
34. Sundararajan,A., Michaud,W.A., Qian,Q., Stahl,G. and Farabaugh,P.J. (1999) Near-cognate peptidyl-tRNAs promote +1 programmed translational frameshifting in yeast. *Mol. Cell*, **4**, 1005–1015.
35. Farabaugh,P.J. (2010) In Atkins,J.F. and Gesteland,R.F. (eds), *Recoding: Expansion of Decoding Rules Enriches Gene Expression*. Springer, New York and Heidelberg, pp. 221–247.
36. Kuchino,Y., Yabusaki,Y., Mori,F. and Nishimura,S. (1984) Nucleotide sequences of three proline tRNAs from *Salmonella typhimurium*. *Nucleic Acids Res.*, **12**, 1559–1562.
37. Björk,G.R. (1996) In Neidhardt,F.C., Curtiss,R.I., Ingraham,J.L., Lin,E.C.C., Low,K.B., Magasanik,B., Resnikoff,W.S., Riley,M., Schaechter,M. and Umbarger,H.E. (eds), *Escherichia coli and Salmonella: Cellular and Molecular Biology*, 2nd edn. ASM press, Washington D.C., pp. 861–886.
38. Näsvall,S.J., Nilsson,K. and Björk,G.R. (2009) The ribosomal grip of the peptidyl-tRNA is critical for reading frame maintenance. *J. Mol. Biol.*, **385**, 350–367.
39. Qian,Q., Li,J.N., Zhao,H., Hagervall,T.G., Farabaugh,P.J. and Björk,G.R. (1998) A new model for phenotypic suppression of frameshift mutations by mutant tRNAs. *Mol. Cell*, **1**, 471–482.
40. Masucci,J.P., Gallant,J., Lindsley,D. and Atkinson,J. (2002) Influence of the *relA* gene on ribosome frameshifting. *Mol. Genet. Genomics*, **268**, 81–86.
41. Pittard,J. (1996) In Neidhardt,F.C., Curtiss,R.I., Ingraham,J.L., Lin,E.C.C., Low,K.B., Magasanik,B., Resnikoff,W.S., Riley,M., Schaechter,M. and Umbarger,H.E. (eds), *Escherichia coli and Salmonella: Cellular and Molecular Biology*, 2nd edn. ASM press, Washington D.C., pp. 458–484.
42. Gavini,N. and Pulakat,L. (1991) Role of translation of the *pheA* leader peptide coding region in attenuation regulation of the *Escherichia coli pheA* gene. *J. Bacteriol.*, **173**, 4904–4907.
43. Hudson,G.S. and Davidson,B.E. (1984) Nucleotide sequence and transcription of the phenylalanine and tyrosine operons of *Escherichia coli* K12. *J. Mol. Biol.*, **180**, 1023–1051.
44. Landick,R., Turnbough,C.L. Jr and Yanofsky,C. (1996) In Neidhardt,F.C., Curtiss,R.I., Ingraham,J.L., Lin,E.C.C., Low,K.B., Magasanik,B., Resnikoff,W.S., Riley,M., Schaechter,M. and Umbarger,H.E. (eds), *Escherichia coli and Salmonella: Cellular and Molecular Biology*, 2nd edn. ASM press, Washington D.C., pp. 1263–1286.
45. Gavini,N. and Pulakat,L. (1991) Role of ribosome release in the basal level of expression of the *Escherichia coli* gene *pheA*. *J. Gen. Microbiol.*, **137**, 679–684.
46. Atkins,J.F., Weiss,R.B. and Gesteland,R.F. (1990) Ribosome gymnastics—degree of difficulty 9.5, style 10.0. *Cell*, **62**, 413–423.
47. Dayhuff,T.J., Atkins,J.F. and Gesteland,R.F. (1986) Characterization of ribosomal frameshift events by protein sequence analysis. *J. Biol. Chem.*, **261**, 7491–7500.
48. Wills,N.M., Ingram,J.A., Gesteland,R.F. and Atkins,J.F. (1997) Reported translational bypass in a *trpR'-lacZ'* fusion is accounted for by unusual initiation and +1 frameshifting. *J. Mol. Biol.*, **271**, 491–498.
49. Haldimann,A. and Wanner,B.L. (2001) Conditional-replication, integration, excision, and retrieval plasmid-host systems for gene structure-function studies of bacteria. *J. Bacteriol.*, **183**, 6384–6393.
50. Neidhardt,F.C., Bloch,P.L. and Smith,D.F. (1974) Culture medium for enterobacteria. *J. Bacteriol.*, **119**, 736–747.

51. Bertani,G. (1951) Studies on lysogenesis. I. The mode of phage liberation by lysogenic *Escherichia coli*. *J. Bacteriol.*, **62**, 293–300.
52. Davis,R.W., Botstein,D. and Roth,J.R. (1980) *A Manual for Genetic Engineering: Advanced Bacterial Genetics*. Cold Spring Harbor Laboratory, New York.
53. Schmieger,H. and Buch,U. (1975) Appearance of transducing particles and the fate of host DNA after infection of *Salmonella typhimurium* with P22-mutants with increased transducing ability (HT-mutants). *Mol. Gen. Genet.*, **140**, 111–122.
54. Scott,J.F., Roth,J.R. and Artz,S.W. (1975) Regulation of histidine operon does not require *hisG* enzyme. *Proc. Natl Acad. Sci. USA*, **72**, 5021–5025.
55. Chen,P., Qian,Q., Zhang,S., Isaksson,L.A. and Björk,G.R. (2002) A cytosolic tRNA with an unmodified adenosine in the wobble position reads a codon ending with the non-complementary nucleoside cytidine. *J. Mol. Biol.*, **317**, 481–492.
56. Näsvall,S.J., Chen,P. and Björk,G.R. (2004) The modified wobble nucleoside uridine-5-oxyacetic acid in tRNAPro(cmo5UGG) promotes reading of all four proline codons *in vivo*. *RNA*, **10**, 1662–1673.
57. Datsenko,K.A. and Wanner,B.L. (2000) One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. *Proc. Natl Acad. Sci. USA*, **97**, 6640–6645.
58. Miller,J.H. (1972) *Experiments in Molecular Genetics*. Cold Spring Harbor Laboratory, New York.
59. Vitreschak,A.G., Lyubetskaya,E.V., Shirshin,M.A., Gelfand,M.S. and Lyubetsky,V.A. (2004) Attenuation regulation of amino acid biosynthetic operons in proteobacteria: comparative genomics analysis. *FEMS Microbiol. Lett.*, **234**, 357–370.
60. Hauben,L., Moore,E.R., Vauterin,L., Steenackers,M., Mergaert,J., Verdonck,L. and Swings,J. (1998) Phylogenetic position of phytopathogens within the Enterobacteriaceae. *Syst. Appl. Microbiol.*, **21**, 384–397.
61. Weiss,R.B., Huang,W.M. and Dunn,D.M. (1990) A nascent peptide is required for ribosomal bypass of the coding gap in bacteriophage T4 gene *60*. *Cell*, **62**, 117–126.
62. Herr,A.J., Gesteland,R.F. and Atkins,J.F. (2000) One protein from two open reading frames: mechanism of a 50 nt translational bypass. *EMBO J.*, **19**, 2671–2680.
63. Wills,N.M., O'Connor,M., Nelson,C.C., Rettberg,C.C., Huang,W.M., Gesteland,R.F. and Atkins,J.F. (2008) Translational bypassing without peptidyl-tRNA anticodon scanning of coding gap mRNA. *EMBO J.*, **27**, 2533–2544.
64. Herr,A.J., Wills,N.M., Nelson,C.C., Gesteland,R.F. and Atkins,J.F. (2004) Factors that influence selection of coding resumption sites in translational bypassing: minimal conventional peptidyl-tRNA:mRNA pairing can suffice. *J. Biol. Chem.*, **279**, 11081–11087.
65. Qian,Q. and Björk,G.R. (1997) Structural alterations far from the anticodon of the tRNAProGGG of *Salmonella typhimurium* induce +1 frameshifting at the peptidyl-site. *J. Mol. Biol.*, **273**, 978–992.
66. Brown,K.D. and Somerville,R.L. (1971) Repression of aromatic amino acid biosynthesis in *Escherichia coli* K-12. *J. Bacteriol.*, **108**, 386–399.
67. Im,S.W. and Pittard,J. (1971) Phenylalanine biosynthesis in *Escherichia coli* K-12: mutants derepressed for chorismate mutase P-prephenate dehydratase. *J. Bacteriol.*, **106**, 784–790.
68. Gowrishankar,J. and Pittard,J. (1982) Regulation of phenylalanine biosynthesis in *Escherichia coli* K-12: control of transcription of the *pheA* operon. *J. Bacteriol.*, **150**, 1130–1137.
69. Gavini,N. and Davidson,B.E. (1990) The *pheR* gene of *Escherichia coli* encodes tRNA(Phe), not a repressor protein. *J. Biol. Chem.*, **265**, 21527–21531.
70. Gollnick,P. and Yanofsky,C. (1990) tRNA(Trp) translation of leader peptide codon 12 and other factors that regulate expression of the tryptophanase operon. *J. Bacteriol.*, **172**, 3100–3107.
71. Keiler,K.C., Waller,P.R. and Sauer,R.T. (1996) Role of a peptide tagging system in degradation of proteins synthesized from damaged messenger RNA. *Science*, **271**, 990–993.
72. Roche,E.D. and Sauer,R.T. (2001) Identification of endogenous SsrA-tagged proteins reveals tagging at positions corresponding to stop codons. *J. Biol. Chem.*, **276**, 28509–28515.
73. Champion,K. and Higgins,N.P. (2007) Growth rate toxicity phenotypes and homeostatic supercoil control differentiate *Escherichia coli* from *Salmonella enterica* serovar Typhimurium. *J. Bacteriol.*, **189**, 5839–5849.
74. Figueroa-Bossi,N., Guerin,M., Rahmouni,R., Leng,M. and Bossi,L. (1998) The supercoiling sensitivity of a bacterial tRNA promoter parallels its responsiveness to stringent control. *EMBO J.*, **17**, 2359–2367.
75. Travers,A. and Muskhelishvili,G. (2005) DNA supercoiling - a global transcriptional regulator for enterobacterial growth? *Nat. Rev Microbiol.*, **3**, 157–169.
76. Mottagui-Tabar,S. and Isaksson,L.A. (1998) The influence of the 5' codon context on translation termination in *Bacillus subtilis* and *Escherichia coli* is similar but different from *Salmonella typhimurium*. *Gene*, **212**, 189–196.
77. Dinçbas-Renqvist,V., Engstrom,A., Mora,L., Heurgue-Hamard,V., Buckingham,R. and Ehrenberg,M. (2000) A post-translational modification in the GGQ motif of RF2 from *Escherichia coli* stimulates termination of translation. *EMBO J.*, **19**, 6900–6907.
78. Mora,L., Heurgue-Hamard,V., de Zamaroczy,M., Kervestin,S. and Buckingham,R.H. (2007) Methylation of bacterial release factors RF1 and RF2 is required for normal translation termination *in vivo*. *J. Biol. Chem.*, **282**, 35638–35645.
79. Baranov,P.V., Gesteland,R.F. and Atkins,J.F. (2002) Release factor 2 frameshifting sites in different bacteria. *EMBO Rep.*, **3**, 373–377.
80. Baranov,P.V. and Gurvich,O.L. (2010) In Atkins,J.F. and Gesteland,R.F. (eds), *Recoding: Expansion of Decoding Rules Enriches Gene Expression*. Springer, New York and Heidelberg, pp. 301–320.